

A COMPARISON OF TEXTURE AND AMPLITUDE BASED UNSUPERVISED SAR IMAGE CLASSIFICATIONS FOR URBAN AREA EXTRACTION

Koray Kayabol and Josiane Zerubia

Ayin Research Team, INRIA Sophia Antipolis Mediterranee,
2004 route des Lucioles, BP93, 06902 Sophia Antipolis Cedex, France

ABSTRACT

We compare the performance of the texture and the amplitude based mixture density models for urban area extraction from high resolution Synthetic Aperture Radar (SAR) images. We use an Auto-Regressive (AR) model with t -distribution error for the textures and a Nakagami density for the amplitudes. We exploit a Multinomial Logistic (MnL) latent class label model as a mixture density to obtain spatially smooth class segments. We combine the Classification EM (CEM) algorithm with the hierarchical agglomeration strategy and a model order selection criterion called Integrated Completed Likelihood (ICL). We test our algorithm on TerraSAR-X data provided by DLR/DFD.

Index Terms— High resolution SAR, classification, texture, multinomial logistic, Classification EM

1. INTRODUCTION

In remote sensing, image classification finds many application areas varying from crop and forest classification to urban area extraction. Finite Mixture Model (FMM) is a suitable statistical model to represent SAR image histogram and to perform a model based classification [1], [2]. The EM algorithm [3] has been used for parameter estimation in latent variable models such as FMM. In this study, we prefer to use an EM based algorithm called Classification EM (CEM) [4], whose computational cost is lower than both the stochastic methods and the conventional EM algorithm. Two drawbacks of the FMM based classification approach using EM algorithm can be sorted as 1) determination of the necessary number of class to represent the data and 2) initialization of the classes [5], [6]. In order to deal with these drawbacks, we combine the hierarchical agglomeration, CEM and ICL [7] criterion as in [8] to obtain an unsupervised classification algorithm which is able to find the necessary number of classes in the mixture model.

In this paper, rather than pixel-based mixture model, we use a block-based FMM which assembles both the SAR amplitudes and the texture statistics into a FMM simultaneously. In this approach, we factorize the block density using the Bayes rule in two parts which are 1) the amplitude density

based on the central pixel of the block and 2) texture density based on the conditional density of the surrounding pixels given the central pixel.

We use a non-Gaussian 2D AR model for residual texture representation. In this autoregressive model, we express a pixel as a linear combination of its neighboring pixels. We assume that the regression error is an independent and identically distributed (iid) Student's t -distribution. t -distribution is a convenient model for robust regression and it has been used in image processing as a robust statistical model [9]. For amplitude based classification, we use the Nakagami density which is a theoretical multi-look SAR amplitude model [2].

The secondary target in land cover classification from SAR images is to find spatially connected and smooth class label maps. A Bayesian approach allows us to include smoothing constraints to classification problems. In our spatial smoothness model, we assign a binary latent class map for each class which indicates the pixels belonging to that class. We introduce the spatial interaction within each binary map adopting multinomial logistic model.

In Section 2 and 3, the mixture model and unsupervised CEM algorithm are given. The simulation results are shown in Section 4. Section 5 presents the conclusion and future work.

2. MULTINOMIAL LOGISTIC MIXTURE OF TEXTURE AND AMPLITUDE BASED DENSITIES

We assume that the observed amplitude $s_n \in \mathbb{R}^+$ at the n th pixel, where $n \in \mathcal{R} = \{1, 2, \dots, N\}$ represents the lexicographically ordered pixel index, is free from any noise and instrumental degradation. We denote \mathbf{s} to be the vector representation of the entire image and \mathbf{s}_n to be the vector representation of the $d \times d$ image block located at n th pixel. Every pixel in the image has a latent class label. Denoting by K the number of classes, we encode the class label as a K dimensional categorical vector \mathbf{z}_n whose elements $z_{n,k}$, $k \in \{1, 2, \dots, K\}$ have the following properties: 1) $z_{n,k} \in \{0, 1\}$ and 2) $\sum_{k=1}^K z_{n,k} = 1$. We may write the probability of \mathbf{s}_n as the marginalization of the joint probability

density $p(\mathbf{s}_n, \mathbf{z}_n | \Theta, \boldsymbol{\pi}_n) = p(\mathbf{s}_n | \mathbf{z}_n, \Theta) p(\mathbf{z}_n | \boldsymbol{\pi}_n)$, [1], as

$$p(\mathbf{s}_n | \Theta, \boldsymbol{\pi}_n) = \sum_{\mathbf{z}_n} \prod_{k=1}^K [p(\mathbf{s}_n | \theta_k) \pi_{n,k}]^{z_{n,k}} \quad (1)$$

where $\boldsymbol{\pi}_n = \{\pi_{n,1}, \dots, \pi_{n,K}\}$ represent the mixture proportions and ensure that $\sum_{k=0}^K \pi_{n,k} = 1$. θ_k are the parameters of the class densities and $\Theta = \{\theta_1, \dots, \theta_K\}$ is the set of the parameters.

Our aim is to use the amplitude and the texture statistics together to classify the SAR images. We may write the density of an image block as a joint density of the central pixel and the surrounding pixels as $p(\mathbf{s}_n | \theta_k) = p(s_n, \mathbf{s}_{\partial n} | \theta_k)$. Using Bayes rule, we factorize the density of the image block as

$$p(\mathbf{s}_n | \theta_k) = p_A(s_n | \theta_k) p_T(\mathbf{s}_{\partial n} | s_n, \theta_k) \quad (2)$$

In this last expression, the first and the second terms represent the amplitude and the texture densities, respectively.

We introduce a t -MRF texture model to use the contextual information for classification. We write the texture model using the neighbors of the pixel in $\mathcal{N}(n)$ as

$$s_n = \sum_{n' \in \mathcal{N}(n)} \alpha_{k,n'} s_{n'} + t_{k,n} \quad (3)$$

where $\alpha_{k,n'}$ are the auto-regression coefficients and the regression errors $t_{k,n}$ are an iid t -distributed zero-mean random variables with β_k degrees of freedom and scale parameters δ_k . In this way, we write the class texture density as a t -distribution such that

$$p_T(\mathbf{s}_{\partial n} | s_n, \boldsymbol{\alpha}_k, \beta_k, \delta_k) = \frac{\Gamma((1 + \beta_k)/2)}{\Gamma(\beta_k/2) (\pi \beta_k \delta_k)^{1/2}} \times \left[1 + \frac{(s_n - \mathbf{s}_{\partial n}^T \boldsymbol{\alpha}_k)^2}{\beta_k \delta_k} \right]^{-\frac{\beta_k+1}{2}} \quad (4)$$

where the vector $\boldsymbol{\alpha}_k$ contains the regression coefficients $\alpha_{k,n'}$.

We model the class amplitudes using Nakagami density, which is a basic theoretical multi-look amplitude model for SAR images [2].

We are able to introduce spatial interactions of the categorical random field by defining a binary spatial auto-regression model. The related probability density of this model is a Multinomial Logistic density which is written as

$$p(\mathbf{z}_n | \mathbf{Z}_{\partial n}, \eta) = \prod_{k=1}^K \left(\frac{\exp(\eta v_k(z_{n,k}))}{\sum_{j=1}^K \exp(\eta v_j(z_{n,j}))} \right)^{z_{n,k}} \quad (5)$$

where $v_k(z_{n,k}) = 1 + \sum_{m \in \mathcal{M}(n), m \neq n} z_{m,k}$ and $\mathbf{Z}_{\partial n} = \{\mathbf{z}_m : m \in \mathcal{M}(n), m \neq n\}$ is the set which contains the neighbors of \mathbf{z}_n in a window $\mathcal{M}(n)$ defined around n . The function $v_k(z_{n,k})$ returns the number of labels which belong to class k in the given window.

Table 1. Unsupervised Amplitude and Texture density mixtures of MnL with CEM (ATML-CEM).

Initialize the classes for $K = K_{max}$.

WHILE $K \geq K_{min}$, DO

$\eta = c, c \geq 0$

WHILE the number of changes $> N \times 10^{-3}$, DO

E-step: Calculate the posteriors of the class labels

C-step: Classify the pixels regarding to the posteriors

M-step: Estimate the parameters of amplitude and texture densities

Update the smoothness parameter η

Find the weakest class

Find the closest class to the weakest class

Merge these two classes

$K \leftarrow K - 1$

3. UNSUPERVISED CLASSIFICATION ALGORITHM

Our strategy follows the same general philosophy as the one proposed in [8], [10]. We start the CEM algorithm with a large number of classes, $K = K_{max}$, and then we reduce the number of classes to $K \leftarrow K - 1$ by merging the weakest class in probability to the one that is most similar to it with respect to a distance measure. The weakest class may be found using the average probabilities of each class. We use a symmetric KL type distance measure called Jensen-Shannon divergence [11] which is defined between two probability density functions to find the closest class to the weakest class. We merge them to constitute a new class and repeat this procedure until we reach the predefined minimum number of classes K_{min} . We determine the necessary number of classes by observing the ICL criterion given in [10]. The summary of the algorithm can be found in Table 1.

4. SIMULATION RESULTS

This section presents the high resolution SAR image classification results of the proposed method called ATML-CEM (Amplitude and Texture density mixtures of MnL with CEM), compared to the corresponding results obtained with other methods. The competitors are Multiphase Level Set (MLS) [12], [13] and K-MnL. We have also tested two different versions of ATML-CEM method which are amplitude based AML-CEM [10] and texture based TML-CEM.

The sizes of the windows for texture and label models are selected to be 3×3 and 13×13 respectively by trial and er-

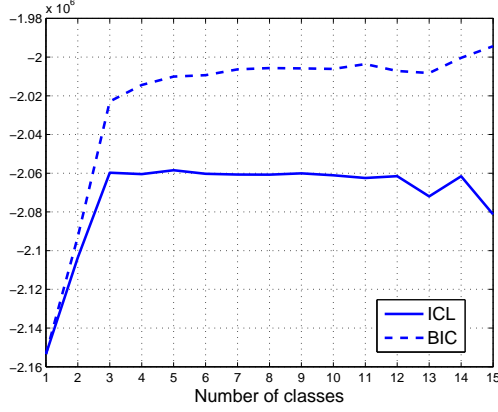


Fig. 2. ICL and BIC values of the classified TSX image for several numbers of classes (from 1 to 15).

ror. We initialize the algorithm as described in [10] and estimate all the parameters along the iterations. MLS method is based on the piecewise constant multiphase Chan-Vese model and implemented by [12], [13]. In this method, we set the smoothness parameter to 2000 and step size to 0.0002 for all data sets. We tune the number of iterations to reach the best result. The K-MnL method is the sequential combination of K-means clustering for classification and Multinomial Logistic label model for segmentation to obtain a fairer comparison with the K-means clustering since K-means does not provide any segmented map.

We tested the algorithms on the following TerraSAR-X image TSX: 900×600 pixels, HH polarized, TerraSAR-X SpotLight (8.2 m ground resolution) 4-look image which was acquired over the city of Rosenheim in Germany (see Fig. 1).

Fig. 3 shows several classification maps found by ATML-CEM with different numbers of classes. From this figure, we can see the evolution of the class maps along the agglomeration based algorithm. We can see the plotted ICL and BIC values with respect to number of classes in Fig. 2. The variations in the ICL and BIC plots are slowed down after 3 or 4 respectively. Since the difference between the values at 3 and 4 is very small and our aim is to find the minimum number of classes, we may say that the mixture model with 3 number of classes is almost enough to represent this data set.

For TSX image in Fig.1, the ground-truth map has been generated manually and covers 20% of the whole image. Fig.1 shows the classification results. The numerical accuracy results are given in Table 2 for 3-classes. In both semi-supervised and unsupervised cases, ATML-CEM provides better results than the others in average. The combination of the amplitude and the texture features helps to increase the quality of classification in average. From Fig. 1, we can see that the MLS and K-MnL methods fail to classify the urban areas. MLS provides a noisy classification map.

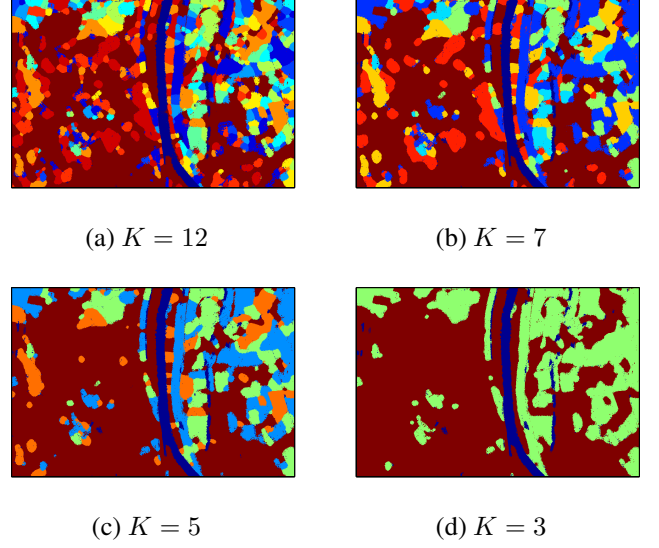


Fig. 3. Classification maps of TSX image obtained with unsupervised ATML-CEM method for different numbers of classes $K = \{3, 5, 7, 12\}$.

Table 2. Accuracy (in %) of the semi-supervised (Ss) and unsupervised (U) classification of TSX image in water, urban and land areas and overall.

	water	urban	land	average
K-MnL (Ss)	100.00	79.03	80.33	86.45
MLS (Ss)	89.47	35.62	84.71	69.93
AML-CEM (U)	92.36	98.29	80.97	90.54
TML-CEM (U)	89.88	96.18	72.32	86.12
ATML-CEM (U)	94.17	98.76	80.93	91.29

5. CONCLUSION

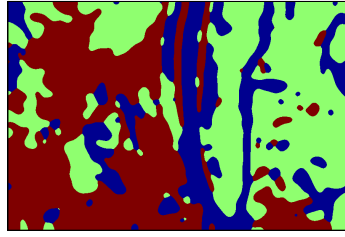
We have proposed a Bayesian model which uses amplitude and texture features together in a FMM along with nonstationary latent class labels. Using these two features together in the model, we obtain better high resolution SAR image classification results for the given SAR image, especially in the urban areas.

6. ACKNOWLEDGMENT

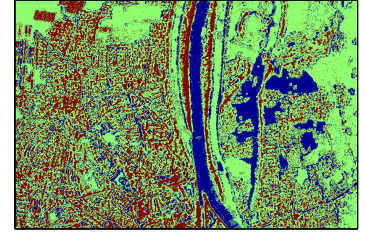
The authors would like to thank Aurélie Voisin (Ayin INRIA, France) for providing the groundtruth and several interesting discussions and Ismail Ben Ayed for providing MLS algorithm [12], [13] online. The first author would like to thank the ERCIM "Alain Bensoussan" Postdoctoral Fellowship Programme for supporting his research at INRIA, France and CWI, Netherlands. The TerraSAR-X image of Rosenheim (©Infoterra) was obtained from <http://www.infoterra.de/>.



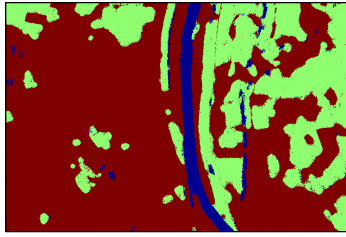
(a) TSX image



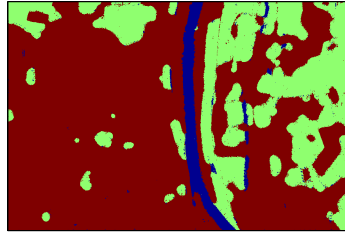
(b) Semi-supervised K-MnL



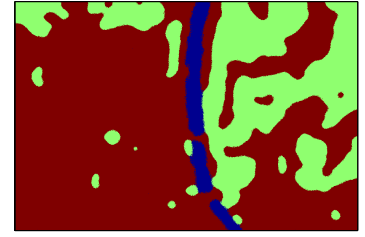
(c) Semi-supervised MLS



(d) Unsupervised ATML-CEM



(e) Unsupervised AML-CEM



(f) Unsupervised TML-CEM

Fig. 1. (a) TSX image and classification maps obtained by (b) semi-supervised K-MnL, (c) semi-supervised MLS, (d) unsupervised ATML-CEM, (e) unsupervised AML-CEM and (f) unsupervised TML-CEM methods. Blue, red and green colors represent water, urban and land areas, respectively.

7. REFERENCES

- [1] D. Titterton, A. Smith, and A. Makov, *Statistical Analysis of Finite Mixture Distributions*, John Wiley & Sons, Chichester (U.K.), 1992.
- [2] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Norwood, 1998.
- [3] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc. B*, vol. 39, pp. 1–22, 1977.
- [4] G. Celeux and G. Govaert, "A classification EM algorithm for clustering and two stochastic versions," *Comput. Statist. Data Anal.*, vol. 14, pp. 315–332, 1992.
- [5] G. Palubinskas, X. Descombes, and F. Kruggel, "An unsupervised clustering method using the entropy minimization," in *Int. Conf. Pattern Recognition, ICPR'98*, 1998, pp. 1816–1818.
- [6] M.A.T. Figueiredo and A.K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 24, no. 3, pp. 381–396, 2002.
- [7] C. Biernacki, G. Celeux, and G. Govaert, "Assessing a mixture model for clustering with the integrated completed likelihood," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 22, no. 7, pp. 719–725, 2000.
- [8] C. Fraley and A. Raftery, "Model-based clustering, discriminant analysis, and density estimation," *J. Am. Statistical Assoc.*, vol. 97, no. 458, pp. 611–631, 2002.
- [9] K. Kayabol, E.E. Kuruoglu, J.L. Sanz, B. Sankur, E. Salerno, and D. Herranz, "Adaptive Langevin sampler for separation of t-distribution modelled astrophysical maps," *IEEE Trans. Image Processing*, vol. 19, no. 9, pp. 2357–2368, 2010.
- [10] K. Kayabol, V.A. Krylov, and J. Zerubia, "Unsupervised classification of SAR images using hierarchical agglomeration and EM," in *MUSCLE Int. Workshop Computational Intelligence for Multimedia Understanding*, 7252, Ed., Pisa, Italy, Dec. 2011, LNCS, pp. 54–65, Springer.
- [11] J. Lin, "Divergence measures based on the Shannon entropy," *J. Am. Statistical Assoc.*, vol. 37, no. 1, pp. 145–151, 1991.
- [12] C. Vazquez, A. Mitiche, and I. Ben Ayed, "Image segmentation as regularized clustering: a fully global curve evolution method," in *Int. Conf. Image Process. ICIIP'04*, Singapore, Oct. 2004, pp. 3467–3470.
- [13] A. Mitiche and I. Ben Ayed, *Variational and Level Set Methods in Image Segmentation*, Springer, Heilderberg, 2010.